SECOND ORDER CHARACTERIZATIONS OF PSEUDOCONVEX FUNCTIONS

by

Mordecai Avriel[1)]

and

Siegfried Schaible[2)]

Technical Report 76-12

June 1976

Department of Operations Research
Stanford University
Stanford, California

SECOND ORDER CHARACTERIZATIONS OF PSEUDOCONVEX FUNCTIONS

by

Mordecai Avriel[1)]

and

Siegfried Schaible[2)]

Technical Report 76-12

June 1976

Department of Operations Research
Stanford University
Stanford, California

1) Technion-Israel Institute of Technology, Haifa, Israel.

2) Industrieseminar, Universität Koln, Koln, West Germany. Research of this author was partially supported by Deutsche Forschungsgemeinschaft, West Germany.

Research and reproduction of this report were partially supported by the Office of Naval Research Contract N00014-75-C-0267; National Science Foundation Grant MCS71-03341 A04; and Energy Research and Development Administration Contract E(04-3)-326 PA #18.

1. Introduction

Convexity plays a central role in the analysis of mathematical programming problems. Numerous generalizations of convex functions have been derived which proved to be useful for extending optimality conditions, previously restricted to convex programs, to larger classes of optimization problems. Some global convergence results of nonlinear programming algorithms can be also extended from convex programs to problems involving certain generalized convex functions. For a review of generalized convexity and its application to mathematical programming see [4, 16, 20, 22, 23]. In this work we shall investigate twice continuously differentiable pseudoconvex functions.

Definition 1. *A real differentiable function* f, *defined on an open convex subset* C *of* R^n *is called pseudoconvex (pcx) if*

$$(x' - x)^T \nabla f(x) \geq 0 \Rightarrow f(x') \geq f(x) \tag{1.1}$$

for all $x \in C$, $x' \in C$. *It is called strictly pseudoconvex (strictly pcx) if*

$$(x' - x)^T \nabla f(x) \geq 0 \Rightarrow f(x') > f(x) \tag{1.2}$$

for all $x \in C$, $x' \in C$, $x \neq x'$.

Pcx and strictly pcx functions generalize convex and strictly convex functions, respectively. It is well known [4] that every local minimum of a pcx function is global and the Kuhn-Tucker necessary conditions are also sufficient for a local (global) minimum in a nonlinear program whose objective function is pcx and the constraints are defined by quasiconvex functions. Furthermore, if the objective function is strictly pcx, there exists at most one global minimum. Global convergence to a minimum by certain numerical algorithms, such as the conjugate gradient method, is ensured in case of pcx functions.

Characterizations of twice differentiable pseudoconvex functions f, in terms of extended Hessians, defined by

$$H(x;\ r(x)) = \nabla^2 f(x) + r(x)\ \nabla f(x)\ \nabla f(x)^T$$

were studied in [2, 4, 5, 6, 21]. In Section 2 of this paper we relate the criteria used in the characterizations to each other and derive additional results of this type. A related topic, discussed in Section 3 is the characterization of pcx functions in terms of bordered determinants. First results in this direction were presented in [1, 10, 12]. Finally, in Section 4, we focus on quadratic functions. A characterization in terms of an extended Hessian [27] is used to develop a necessary and sufficient condition for (strictly) pcx quadratic functions in terms of bordered determinants.

2. Pseudoconvexity in terms of extended Hessians

Throughout this paper we shall always refer to f as a twice continuously differentiable real function, defined on an open convex subset C of R^n. Vectors are considered to be column vectors.

Accordingly, if x is a vector, its transpose is denoted by x^T. Let Z denote the set of normalized direction vectors in R^n, that is

$$Z = \{z \in R^n : \|z\| = 1\} .$$

Proposition 1. *If f is pseudoconvex on C then there exists a function $\rho : CxZ \to R$ such that*

$$z^T[\nabla^2 f(x) + \rho(x,z)\ \nabla f(x)\ \nabla f(x)^T]z \geq 0 \tag{2.1}$$

for all $x \in C$, $z \in Z$.

Proof. First we note that for a pcx function f we have $z^T \nabla^2 f(x) z \geq 0$ if $z^T \nabla f(x) = 0$ (see Lemma 6.2 in [2]). Consider now

$$\rho_0(x,z) = \begin{cases} 0 & \text{if } z^T\nabla f(x) = 0 \\ \\ -\dfrac{z^T\ \nabla^2 f(x) z}{[z^T\ \nabla f(x)]^2} & \text{if } z^T\nabla f(x) \neq 0 . \end{cases} \tag{2.2}$$

It follows that $\rho=\rho_0$ satisfies (2.1) as asserted. □

Note that if ρ is any other function satisfying (2.1), then

$$\rho(x,z) \geq \rho_0(x,z)$$

for all $x \in C$, $z \in Z$ such that $z^T\nabla f(x) \neq 0$. Define now

$$r(x) = \sup\{\rho(x,z) : z \in Z\} . \tag{2.3}$$

If $r_0(x) = \sup\{\rho_0(x,z): z \in Z\}$ is finite on C, then there exists a function r, depending on x, such that

$$H(x; r(x)) = \nabla^2 f(x) + r(x) \nabla f(x) \nabla f(x)^T \tag{2.4}$$

is positive semidefinite on C, and conversely. Furthermore, f is r_0-convex on C [2, 4, 17, 20] if and only if

$$r_0 = \sup\{r_0(x): x \in C\}$$

is finite. For many pcx functions r_0 is unbounded on C, as in the case of quadratic functions on maximal sets of pseudoconvexity [27]. Example 1 below demonstrates a case of a pseudoconvex function for which there exists no finite $r_0(x)$ for any $x \in C$. Hence, positive semidefiniteness of the extended Hessian given by (2.4) is not a necessary condition for pseudoconvexity, as it was erroneously stated in [21]. Consequently, the characterization of pcx functions in [6] applies only to those functions where $H(x; r(x))$ is positive semi-definite for some $r(x)$.

Example 1. Let $f(x) = x_2/x_1$ on $C = \{x \in R^2: x_1 > 0\}$. Then f is pcx on C [16]. For any $x \in C$ and $z \in Z$ such that $z_1 \neq 0$, $z_2/z_1 \neq x_2/x_1$ we have

$$\rho_0(x,z) = 2[(z_2/z_1) - (x_2/x_1)]^{-1} . \tag{2.5}$$

Since $\rho_0(x,z) \to +\infty$ as $(z_2/z_1) \searrow (x_2/x_1)$, it follows that $r_0(x) = +\infty$. ||

The condition stated in Proposition 1 is necessary for pseudo-convexity. However, it is not sufficient, as can be seen in the following example.

Example 2. Let $f(x) = (x)^3$, $C = R$. Here

$$\rho_0(x,z) = \begin{cases} 0 & \text{if } x = 0 \\ -\frac{2}{3}(x)^{-3} & \text{if } x \neq 0 . \end{cases}$$

Clearly, inequality (2.1) holds, but f is not pcx. ||

We observe that in the last example ρ_0 is unbounded on intervals containing the origin. However, ρ_0 is (locally) bounded on open intervals that do not contain the origin. On those intervals f happens to be pcx. This example motivates our next result. For $x \in C$ and $z \in Z$, let

$$T(x,z) = \{t \in R: x + tz \in C\} .$$

Then we have

Proposition 2. If there exists a $\rho : C \times Z \to R$ *satisfying* (2.1) *such that for every* $x \in C$, $z \in Z$ *and every compact interval* $I \subset T(x,z)$ *the number*

$$\omega(I) = \sup\{\rho(x + tz,z):t \in I\} \tag{2.6}$$

is finite, then f *is pseudoconvex on* C.

Proof. Define $h(t) = f(x + tz)$ for $t \in I$. From (2.1) we have

$$h''(t) + \rho(x + tz,z)(h'(t))^2 \geq 0$$

for $t \in I$. By (2.6) $\omega(I) < +\infty$ and it follows that

$$h''(t) + \omega(I)\ (h'(t))^2 \geq 0$$

for all $t \in I$. Hence, h is $\omega(I)$-convex (in the sense of r-convexity) on I which implies psuedoconvexity. Therefore, h is pcx on $T(x,z)$, for all $x \in C$ and $z \in Z$. This implies pseudoconvexity of f on C.

$\square$

Proposition 2 shows that f is pcx on C if h is $\omega(I)$-convex on all compact intervals I in $T(x,z)$. We shall see in the sequel that this condition is also necessary, provided certain "pathological" functions (an example of which is given below), are excluded.

Example 3. Let

$$f(x) = \begin{cases} -\int_0^x (\xi)^4 \, [2 + \sin(1/\xi)] \, d\xi & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ \int_0^x (\xi)^4 \, [2 + \sin(1/\xi)] \, d\xi & \text{if } x > 0 \end{cases}$$

This function is strictly pcx on R. To see this, consider

$$f'(x) = \begin{cases} -(x)^4 \, [2 + \sin(1/x)] & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ (x)^4 \, [2 + \sin(1/x)] & \text{if } x > 0 \end{cases}$$

and $f'(x) > 0$ for $x > 0$, $f'(0) = 0$, and $f'(x) < 0$ for $x < 0$. However, by computing the second derivative

$$f''(x) = \begin{cases} -(x)^2 \, \{4x[2 + \sin(1/x)] - \cos(1/x)\} & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ (x)^2 \, \{4x[2 + \sin(1/x)] - \cos(1/x)\} & \text{if } x > 0 \end{cases}$$

we can see that f'' changes sign in every neighborhood of the origin, thus f is not convex. Take now $x = 0$ and $z = 1$. Then, $x + tz = t$ and for $t > 0$ we have

$$\rho_0(t,1) = \frac{-f''(t)}{[f'(t)]^2} = \frac{-\{4t[2+\sin(1/t)] - \cos(1/t)\}}{(t)^6\ [2+\sin(1/t)]^2}$$

For $k = 1,2,\ldots,$ let $t^k = 1/(2k\pi)$. Thus

$$\rho_0(t^k,1) = \frac{1-8t^k}{4(t^k)^6} = 8(2k\pi - 8)(\pi)^5(k)^5$$

and $\rho_0(t^k,1) \to +\infty$ as $k \to \infty$. It follows that the sufficient condition of Proposition 2 is not necessary for this pcx function. ||

In the next proposition we restrict the discussion to functions for which the second order neighborhood sufficient conditions of optimality [4,13] are also necessary. First we need

Definition 2. *A twice differentiable function* $h:R \to R$ *is said to be regular on an open interval* (a,b) *if the following statements are equivalent:*

(i) $h(t)$ *has a local minimum at* $t^* \in (a,b)$

(ii) $h'(t^*) = 0$ *and* $h''(t) \geq 0$ *in some neighborhood* $N(t^*) \subset (a,b)$.

The function $f:C \to R$ *is said to be regular on* C *if* $h(t) = f(x+tz)$ *is regular on* $T(x,z)$ *for all* $x \in C$, $z \in Z$.

The function f appearing in Example 3 is not regular on R.

Proposition 3. Let f be regular on C. Then f is pseudoconvex on C if and only if for every $x \in C$, $z \in Z$ and every compact interval $I \subset T(x,z)$ there exists a number $\omega(I)$ such that $h(t) = f(x + tz)$ is $\omega(I)$-convex on I.

Proof. The sufficiency part has been proven in Proposition 2. To prove necessity, let $x \in C$, $z \in Z$ and consider $h(t) = f(x + tz)$ for $t \in I$. From Proposition 1 we have that

$$h''(t) + \rho_0(x + tz,z)\,[h'(t)]^2 \geq 0 \tag{2.7}$$

for $t \in I$, where

$$\rho_0(x + tz,z) = \begin{cases} 0 & \text{if } h'(t) = 0 \\ \dfrac{-h''(t)}{[h'(t)]^2} & \text{if } h'(t) \neq 0 \end{cases} \tag{2.8}$$

For every $\bar{t} \in I$ such that $h'(\bar{t}) \neq 0$ there exists a neighborhood $N(\bar{t})$ such that $h'(t) \neq 0$ for $t \in N(\bar{t})$. Thus, $\rho_0(x + tz,z)$ is bounded there and can be replaced by some $\tilde{\rho}(\bar{t})$ in $N(\bar{t})$ without violating (2.7).

Every $\bar{t} \in I$ such that $h'(\bar{t}) = 0$ is a local minimum, because h is pcx. Since h is regular, $h''(t) \geq 0$ for every t in some neighborhood $N(\bar{t})$. It follows that $\rho_0(x + tz,z)$ can be replaced by $\tilde{\rho}(\bar{t}) = 0$ in $N(\bar{t})$ without violating (2.7). Since I is compact, a finite number of neighborhoods $N(t^i)$ associated with $t^i \in I$ will cover I. Let $\omega(I) = \sup \tilde{\rho}(t^i)$. Then we have

$$h''(t) + \omega(I)\,[h'(t)]^2 \geq 0 \tag{2.9}$$

for every $t \in I$. □

As we have already seen, there are pcx functions for which

$$r_0(x) = \sup\{\rho_0(x,z) : z \in Z\} \tag{2.10}$$

is not necessarily finite. In the following we consider functions for which $r_0(x)$ is finite for every $x \in C$. The next result follows then from Proposition 2.

Proposition 4. *If there exists a continuous function $r: C \to R$ such that $H(x; r(x))$ is positive semidefinite, then f is pseudoconvex on C.*

It is interesting to relate the preceding results to functions which are convex transformable, that is, they can be transformed into convex functions by a monotone transformation. The family of G-convex functions was introduced and studied in [5]. A function $f: C \to R$ is called G-convex if there exists a twice continuously differentiable function $G: D \to R$, $G'(y) > 0$, such that $Gf: C \to R$ is convex on C, where $D \subset R$ contains the range of f. We have then

Proposition 5 [5]. *The function* $f:C \to R$ *is* G-*convex on* C *if and only if there exists a twice continuously differentiable function* $G:D \to R$, $G'(y) > 0$ *such that*

$$\nabla^2 f(x) + \frac{G''(f(x))}{G'(f(x))} \nabla f(x) \nabla f(x)^T \tag{2.11}$$

is positive semidefinite for all $x \in C$.

It immediately follows from Proposition 5 that G-convex functions satisfy the conditions of Proposition 4.

We introduce now the following notation:

T = family of G-convex functions on C.

H = family of functions for which a positive semidefinite extended Hessian $H(x;r(x))$ exists at every $x \in C$.

H_c = subfamily of H, with a continuous r on C.

P = family of pcx functions on C.

In view of Propositions 4 and 5 we have

$$T \subset H_c \subset P \ . \tag{2.12}$$

Examples 1 and 3 illustrate the fact that $H_c \neq P$. The following example shows that $T \neq H_c$, thus generally the inclusions in (2.12) are strict.

Example 4. Let $C \subset R^2$ be an open convex set contained in

$$M = \{x: -[-\frac{3}{4}(x_1)^3]^{1/2} < x_2 < [-\frac{3}{4}(x_1)^3]^{1/2}\} .$$

The function

$$f(x) = (x_1)^3 + (x_2)^2$$

is pcx on C [23,26]. Moreover, $f \in H_c$ with

$$r_0(x) = \frac{1}{2}\left|\frac{3}{4}(x_1)^3 + (x_2)^2\right|^{-1} .$$

Now assume that there exists a point $\bar{x}$ on the boundary of C which is also on the boundary of M, i.e. $\frac{3}{4}(\bar{x}_1)^3 + (\bar{x}_2)^2 = 0$. Then $r_0(x)$ becomes arbitrarily large approaching $\bar{x}$ from within C. If f were convex transformable, then $H(x;r(x))$ would be positive semidefinite for some $r(x)$ which is constant, and therefore finite, on $\{x \in C: f(x) = f(\bar{x})\}$ as can be seen from (2.11). However, $r_0(x)$ is not finite there. Therefore, $f \notin T$. ||

We may mention here that for quadratic functions the families T, H_c and P are equivalent, that is, $T = H_c = P$ [27].

The following result characterizes T as a subset of H_c.

Proposition 6. *$f \in T$ if and only if $f \in H_c$ and*

$$\alpha(y) = \sup\{r(x): f(x) = y, x \in C\}$$

is finite for all y in the range of f.

Proof. If $f \in T$, then, by Proposition 5, $f \in H_c$ with $r(x) = G''(f(x))/G'(f(x))$. Hence for all y in the range of f, $\alpha(y)$ is finite.

Conversely, let $f \in H_c$ and suppose $\alpha(y)$ is finite for all y in the range of f. Since f and r are continuous, so is α. The differential equation

$$\frac{G''(y)}{G'(y)} = \alpha(y)$$

has a solution

$$G(y) = \int^{y} \exp\left(\int^{\eta} \alpha(\tau)d\tau\right) d\eta$$

with $G'(y) > 0$. Hence $f \in T$. □

From Proposition 6 we immediately obtain

Proposition 7. If $f \in H_c$ *and*

$$\beta(y) = \sup\{r(x) : f(x) \leq y,\ x \in C\}$$

is finite for all y *in the range of* f, *then* $f \in T$.

The converse of the last proposition is not true, as can be seen from

Example 5. Let

$$C = \{x \in R^2 : x_1 > 0,\ x_2 > 0\}$$

and let

$$f(x) = \frac{(x_1)^{3/2}}{x_2}.$$

For $x \in C$, $r_0(x) = 1/f(x)$ is the smallest number such that $H(x;r(x))$ is positive semidefinite. Then $\beta_0(y) = \sup\{1/f(x) : f(x) \le y,\ x \in C)\} = +\infty$ On the other hand, taking $G(y) = y^2$ we can see that $(f(x))^2$ is convex on C [24]. ||

From Proposition 7 we obtain

Proposition 8. If $f \in H_c$ *and the level sets*

$$S(f,y) = \{x \in C,\ f(x) \le y\}$$

are compact for every $y \in R$, *then* $f \in T$.

Compactness of the level sets is, however, not necessary for functions in H_c to be convex transformable. For example, all nonconvex pcx quadratic functions have unbounded level sets, but they are convex transformable [26,27].

We conclude this section by presenting some related results on strictly pcx functions. Here we use the following notation:

T^s = subfamily of T, where $\nabla^2 Gf(x)$ is positive definite on C.

H^s = subfamily of H, where $H(x;r(x))$ is positive definite on C.

H_c^s = subfamily of H^s, where r is continuous on C.

P^s = family of strictly pcx functions on C.

Then, from Propositions 4 and 5 we have

$$T^s \subset H_c^s \subset P^s .$$

Although we have seen that $H_c \neq H$, we now state and prove

Proposition 9. Using the above notation,

$$H_c^s = H^s .$$

Proof. We have to show that $H^s \subset H_c^s$. Let K be a compact set in C and let $\bar{x} \in K$. Since $f \in H^s$, there exists an $\epsilon > 0$ such that $z^T\nabla^2 f(\bar{x})z > 0$ in $Z_1 = \{z \in Z: |z^T\nabla f(\bar{x})| < \epsilon\}$. Because of continuity of ∇f and $\nabla^2 f$ there exists a neighborhood $N(\bar{x})$ of $\bar{x}$ such that for all $x \in N(\bar{x})$ we have $z^T\nabla^2 f(x)z > 0$ in Z_1 and $|z^T\nabla f(x)| \geq \epsilon/2 > 0$ in $Z \setminus Z_1$. Let

$$r(\bar{x}) > \max\{0, \max\{-z^T\nabla^2 f(x)z/(z^T\nabla f(x))^2 : x \in N(\bar{x}),\ z \in Z \setminus Z_1\}\}.$$

The right-hand side is finite, and we see that $z^T H(x;r(\bar{x}))z > 0$ for all $z \in Z$, $x \in N(\bar{x})$. Since K is compact, there exists a finite number of neighborhoods $N(x^k)$, $x^k \in K$, that cover K. Thus $H(x;r(K))$ is positive definite in K for $r(K) = \sup\{r(x^k)\}$.

Now let $\{K_i\}$ be a sequence of compact convex sets in R^n such that $K_i \subset K_{i+1}$, $i = 1,2,\ldots$, and $C = \bigcup_{i=1}^{\infty} K_i$. As we saw before, there exists a number $r(K_i)$ such that $H(x;r(K_i))$ is positive definite for all $x \in K_i$. Since $K_i \subset K_{i+1}$, it can be assumed that $r(K_i) \leq r(K_{i+1})$. Since the K_i are assumed to be compact convex, it follows that a continuous function r on C can be constructed such that $r(x) \geq r(K_1)$ at $x \in K_1$ and $r(x) \geq r(K_i)$ at $x \in K_i \setminus K_{i-1}$ for $i = 2,\ldots$ [28]. Since then $H(x;r(x))$ is positive definite on C, we proved that $f \in H_c^s$. $\square$

It should be noted that the inclusions $T^s \subset T$, $H_c^s \subset H_c$ and $P^s \subset P$ can be strict, as illustrated below.

Example 6. Let $f(x_1,x_2) = \ln x_1$ and let $C = \{x \in R^2 : x_1 > 0\}$. Then $f \in T$, but $f \notin T^s$. $\|$

The results of this section can be summarized in the following schematic representation:

$$\begin{tikzcd}
 & & & H \\
T \arrow[r, Rightarrow, "4"] & H_c \arrow[rru, Rightarrow, "2{,}3"] \arrow[rr, Rightarrow, "1{,}3"] & & P \\
T^s \arrow[u, Rightarrow, "6"'] \arrow[r, Rightarrow, "4"] & H_c^s \arrow[u, Rightarrow, "6"'] \arrow[r, Leftrightarrow] & H^s \arrow[r, Rightarrow, "3"] & P^s \arrow[u, Rightarrow, "6"']
\end{tikzcd}$$

The numbers appearing next to the arrows refer to the example numbers which show that the reverse implications generally do not hold. Families H and P are not related to each other by inclusion, as can be seen from Examples 1 and 2.

3. Pseudoconvexity in terms of bordered determinants.

In this section we shall deal with determinants of the bordered Hessian of f, given by

$$B_f(x) = \begin{pmatrix} 0 & \nabla f(x)^T \\ \nabla f(x) & \nabla^2 f(x) \end{pmatrix}$$

as related to pseudoconvexity. Let Q_k be the set consisting of monotone increasing sequences of k numbers from $\{1,\ldots,n\}$, that is

$$Q_k = \{\gamma:\gamma = (i_1,\ldots,i_k),\ 1 \leq i_1 < \cdots < i_k \leq n\} .$$

Let $H_{\gamma,k}$ denote the principal minor of order k of the $n \times n$ Hessian $\nabla^2 f$, formed by the $i_1,\ldots,i_k$ rows and columns of $\nabla^2 f$. The leading principal minors of $\nabla^2 f$ are denoted by H_k, $k = 1, \ldots, n$. We associate with B_f and $H_{\gamma,k}$ the principal minor

$$D_{\gamma,k} = \det \begin{pmatrix} 0 & \frac{\partial f}{\partial x_{i_1}} & \cdots & \frac{\partial f}{\partial x_{i_k}} \\ \frac{\partial f}{\partial x_{i_1}} & \frac{\partial^2 f}{\partial x_{i_1}^2} & \cdots & \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_k}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f}{\partial x_{i_k}} & \frac{\partial^2 f}{\partial x_{i_k} \partial x_{i_1}} & \cdots & \frac{\partial^2 f}{\partial x_{i_k}^2} \end{pmatrix} .$$

Similarly, D_k will denote the leading principal minor of order k+1 of B_f. We shall refer to $D_{\gamma,k}$ and D_k as <u>bordered determinants</u>. Characterizations of the families of functions introduced in the previous section in terms of bordered determinants will be presented now.

First we need

<u>Proposition 10</u>. <u>Let</u> A <u>be a real</u> $k \times k$ <u>matrix and let</u> $b \in R^k$. <u>Then, for any real number</u> r <u>we have</u>

$$\det(A + rbb^T) = \det A - r \det \begin{pmatrix} 0 & b^T \\ b & A \end{pmatrix} \tag{3.1}$$

<u>Proof</u>. Suppose that $r \neq 0$. Then, for Schur's formula [14] we obtain

$$\det \begin{pmatrix} -\frac{1}{r} & b^T \\ b & A \end{pmatrix} = -\frac{1}{r} \det(A + rbb^T) . \tag{3.2}$$

It is easy to show that

$$\det\begin{pmatrix} -\frac{1}{r} & b^T \\ b & A \end{pmatrix} = -\frac{1}{r}\det A + \det\begin{pmatrix} 0 & b^T \\ b & A \end{pmatrix} \tag{3.3}$$

and (3.1) follows from equating the right-hand sides of (3.2) and (3.3). □

We can state and prove now

Proposition 11. *A function f belongs to the family H if and only if*

$$H_{\gamma,k}(x) - r(x)\, D_{\gamma,k}(x) \geq 0 \tag{3.4}$$

for all $x \in C$ *and* $\gamma \in Q_k$, $k = 1,\dots,n$.

Proof. The family H consists of all functions f for which a positive semidefinite $H(x;r(x))$ exists at every $x \in C$. Since a square matrix is positive semidefinite if and only if all its principal minors are nonnegative, the proof follows from Proposition 10. □

We also have

Proposition 12. *A function f belongs to H if and only if* $D_{\gamma,k}(x) \leq 0$ *and if* $D_{\gamma,k}(x) = 0$ *then* $H_{\gamma,k}(x) \geq 0$ *for all* $x \in C$ *and* $\gamma \in Q_k$, $k = 1,\dots,n$.

Proof. Replace $r(x)$ in (3.4) by any arbitrarily large $\bar{r}(x)$. □

A square matrix is positive definite if and only if all its leading principal minors are positive. Consequently, we have the following analogous result to Proposition 12.

Proposition 13. *A function f belongs to H^s if and only if $D_k(x) \leq 0$ and if $D_k(x) = 0$, then $H_k(x) > 0$ for all $x \in C$ and $k = 1,\ldots,n$.*

We have seen in the previous section that H_c (and not H) consists of pcx functions only. Let us state and prove now a sufficient condition in terms of bordered determinants for a function to belong to H_c, and thus to be pcx.

Proposition 14. *Suppose that $D_{\gamma,k}(x) \leq 0$ for all $x \in C$ and $\gamma \in Q_k$, $k = 1,\ldots,n$, and if $D_{\gamma,k}(\bar{x}) = 0$, then $H_{\gamma,k}(x) \geq 0$ for all x in some neighborhood $N_{\gamma,k}(\bar{x})$ of $\bar{x}$. Then $f \in H_c$, and thus f is pseudo-convex on C.*

Proof. By Proposition 12, we only have to show that a continuous r can be found such that $H(x;r(x))$ is positive semidefinite. As shown in the proof of Proposition 9, it suffices to prove that for all compact sets K in C there exists a number $r(K)$ such that $H(x;r(K))$ is positive semidefinite for $x \in K$. Let, therefore, $K \subset C$ be compact and let $\bar{x} \in K$.

If $D_{\gamma,k}(\bar{x}) < 0$, then $D_{\gamma,k}(x) < 0$ in some neighborhood $N_{\gamma,k}(\bar{x})$. Define

$$r_{\gamma,k}(\bar{x}) = \sup\{(H_{\gamma,k}(x)-1)/D_{\gamma,k}(x) : x \in N_{\gamma,k}(\bar{x})\} .$$

Then $H_{\gamma,k}(x) - r_{\gamma,k}(\bar{x})\, D_{\gamma,k}(x) \geq 1 \geq 0$ for $x \in N_{\gamma,k}(\bar{x})$.

If $D_{\gamma,k}(\bar{x}) = 0$, then $H_{\gamma,k}(x) \geq 0$ in some neighborhood $N_{\gamma,k}(\bar{x})$ by assumption. Thus

$$H_{\gamma,k}(x) - r_{\gamma,k}(\bar{x})\, D_{\gamma,k}(x) \geq H_{\gamma,k}(x) \geq 0$$

for $x \in N_{\gamma,k}(\bar{x})$, where $r_{\gamma,k}(\bar{x}) = 1$.

Let $N(\bar{x}) = \bigcap_{\gamma,k} N_{\gamma,k}(\bar{x})$ and $r(\bar{x}) = \max_{\gamma,k} r_{\gamma,k}(\bar{x})$. Thus for each $\bar{x} \in K$ there exists a neighborhood $N(\bar{x})$ and a number $r(\bar{x})$ such that $H_{\gamma,k}(x) - r(\bar{x})\, D_{\gamma,k}(x) \geq 0$ for all $x \in N(\bar{x})$ and $\gamma \in Q_k$, $k = 1,\ldots,n$. In view of Proposition 11, $H(x;r(\bar{x}))$ is positive semidefinite in $N(\bar{x})$. Since the compact set K is covered by finitely many neighborhoods $N(x^k)$ of points $x^k \in K$, there exists a number $r(K)$ such that $H(x;r(K))$ is positive semidefinite on K. □

Ferland [10,12], extending previous results of Arrow and Enthoven [1], has considered the following families of functions:

$D_<$ = family of functions for which $D_k(x) < 0$ for all $x \in C$ and $k = 1,\ldots,n$,

$D_\leq$ = family of functions for which $D_k(x) \leq 0$ for all $x \in C$ and $k = 1,\ldots,n$.

Using this notation, Ferland proved that

$$D_< \subset P \subset D_\leq \,. \tag{3.5}$$

Let us introduce now two additional families.

D^s = *family of functions for which* $D_k(x) \leq 0$, and if $D_k(x) = 0$, then $H_k(x) > 0$ for all $x \in C$ and $k = 1,\ldots,n$.

D = family of functions for which $D_{\gamma,k}(x) \leq 0$ for all $x \in C$ and $\gamma \in Q_k$, $k = 1,\ldots,n$, and if $D_{\gamma,k}(\bar{x}) = 0$, then $H_{\gamma,k}(x) \geq 0$ for all x in some neighborhood $N_{\gamma,k}(\bar{x})$ of $\bar{x}$.

In Proposition 13 we proved that $D^s = H^s$ and $H^s \subset D$. Hence, we have

<u>Proposition 15</u>.

$$D_< \subset D^s = H^s \subset D \subset D_\leq \ . \tag{3.6}$$

Note that the first inequality in (3.5), that is, $D_< \subset P$ follows from (3.6), since by Proposition 14, $D \subset P$. Also, since $H^s \subset P^s$ (see Section 2), the first inclusion in (3.6) shows that $D_<$ covers only <u>strictly</u> pcx functions. Let us show now that the inclusion $D_< \subset D^s$ can be strict, that is, there are strictly pcx functions in D^s which do not belong to $D_<$.

<u>Example 7</u>. Let $f(x_1,x_2) = \ln[(x_1)^2 + (x_2)^2]$ and $C = \{x \in R^2 : x_2 > 0\}$. For this function

$$D_1(x) = -\frac{4(x_1)^2}{[(x_1)^2 + (x_2)^2]^2} \leq 0 \ .$$

If $D_1(\bar{x}) = 0$, then

$$H_1(\bar{x}) = \frac{-2(\bar{x}_1)^2 + 2(\bar{x}_2)^2}{[(\bar{x}_1)^2 + (\bar{x}_2)^2]^2} = \frac{2}{(\bar{x}_2)^2}$$

which is positive. Furthermore,

$$D_2(x) = -\frac{8}{[(x_1)^2 + (x_2)^2]^2} < 0 \ .$$

Hence, $f \in D^s$ but $f \notin D_<$. ||

Since the inclusion $D^s \subset D$ is strict as we have shown in Example 6, Proposition 14 presents a stronger condition than that derived from Proposition 13; Proposition 14 also covers pcx functions which are not strictly pcx.

Proposition 13 and Proposition 14 yield the strongest sufficient conditions in terms of bordered determinants for strict pseudoconvexity and pseudoconvexity, respectively, known so far. From Proposition 12 we can see that $H \subset D_{\leq}$ and by (3.5) we have $P \subset D_{\leq}$. Both inclusions are strict as can be seen below.

Example 8 [10]. Let $f(x_1, x_2) = -(x_1 + x_2)^2$ and $C = R^2$. Then $f \in D_{\leq}$, but $f \notin P$ and $f \notin H$, since along the line $x_1 + x_2 = 0$ the extended Hessian of f is not positive semidefinite. ||

We now complete the schematic representation of the relations between the various families mentioned above as follows:

$$
\begin{tikzcd}
 & & & H \arrow[drr, Rightarrow, "8"] & & \\
T \arrow[rr, Rightarrow, "4"] & & H_c \arrow[rr, Rightarrow, "1,3"] \arrow[ur, Rightarrow, "2,3"] & & P \arrow[r, Rightarrow, "8"] & D_{\le} \\
 & & D \arrow[u, Rightarrow] & & & \\
T^s \arrow[uu, Rightarrow, "6"] \arrow[r, Rightarrow, "4"] & D^s \arrow[r, Leftrightarrow] & H_c^s \arrow[u, Rightarrow, "6"'] \arrow[r, Leftrightarrow] & H^s \arrow[r, Rightarrow, "3"] & P^s \arrow[uu, Rightarrow, "6"'] & \\
D_< \arrow[ur, Rightarrow, "7"] & & & & &
\end{tikzcd}
$$

The families T^s and $D_<$ are not related to each other as can be seen from Examples 4 and 7. We could not find an example showing that the inclusion $D \subset H_c$ is strict. Finally, we may remark that the sign test of the leading principal minors of an extended or bordered Hessian can be conveniently performed by a procedure suggested by Cottle [7].

4. Pseudoconvexity of quadratic functions.

In this section we focus attention on quadratic functions of the form

$$q(x) = x^T Q x + a^T x \tag{4.1}$$

on an open convex set $C \subset R^n$. Characterizations of pcx quadratic functions were derived in [8-11, 18, 19, 23, 25-27]. In [23,26] it was shown that pcx quadratic functions are G-convex. Using this result, a characterization in terms of $H(x;r(x))$ was presented in [27], generalizing a result in [21]. Restricting the families introduced in the previous sections to families of quadratic functions it was shown in [27] that

$$T = H_c = H = P \quad \text{and} \quad T^s = H_c^s = H^s = P^s. \tag{4.1}$$

Let us characterize now pcx quadratic functions in terms of bordered determinants. We have

Proposition 16. *A quadratic function* q *is pseudoconvex on* C *if and only if* $D_{\gamma,k}(x) \leq 0$, *and if* $D_{\gamma,k}(x) = 0$, *then* $H_{\gamma,k}(x) \geq 0$ *for all* $x \in C$ *and* $\gamma \in Q_k$, $k = 1,\dots,n$.

The proof of this proposition follows from the fact that $H = P$ for quadratic functions and from Proposition 12. Since the $H_{\gamma,k}$ are constant for q, Proposition 16 shows that $D = P$.

For strictly pcx functions we have seen in Proposition 13 that $D^s = H^s$. Since $H^s = P^s$ for quadratic functions, we have

Proposition 17. *A quadratic function* q *is strictly pseudoconvex on* C *if and only if* $D_k(x) \leq 0$, *and if* $D_k(x) = 0$, *then* $H_k(x) > 0$ *for all* $x \in C$ *and* $k = 1,\dots,n$.

The relationship between families of pcx quadratic functions can be represented by the following schematic diagram:

$$\begin{tikzcd}
 & T \arrow[r, Leftrightarrow] & H_c \arrow[r, Leftrightarrow] & H \arrow[r, Leftrightarrow] & D \arrow[r, Leftrightarrow] & P \arrow[r, Rightarrow, "8"] & D_{\leq} \\
D_{<} \arrow[r, Rightarrow, "10"] & T^s \arrow[r, Leftrightarrow] & H_c^s \arrow[r, Leftrightarrow] & H^s \arrow[r, Leftrightarrow] & D^s \arrow[u, Rightarrow, "9"'] \arrow[r, Leftrightarrow] & P^s &
\end{tikzcd}$$

The next two examples respectively demonstrate that the inclusions $D^s \subseteq D$ and $D_< \subseteq T^s$ can be strict.

Example 9. Let $q(x_1,x_2) = -(x_1)^2$ and $C = \{x \in R^2 : x_1 > 0\}$. It is easy to verify that $q \in D$. But $H_k(x) \leq 0$ for all $x \in C$ and $k = 1,2$. Hence $q \notin D^s$. ||

Example 10. Let $q(x_1,x_2) = (x_1)^2 + x_2$ and $C = R^2$. $D_1(x) = -4(x_1)^2$, $D_2(x) = -2 < 0$. If $D_1(\bar{x}) = 0$, then $H_1(\bar{x}) = 2 > 0$. Hence $q \notin D_<$ and $q \in D^s$. ||

In conclusion, the families D and D^s respectively characterize pcx and strictly pcx quadratic functions in terms of bordered determinants.

Acknowledgment. This research was carried out during the authors' visit at the Department of Operations Research of Stanford University.

REFERENCES

[1] K.J. Arrow and A.D. Enthoven, "Quasi-concave programming," Econometrica 29 (1961), 779-800.

[2] M. Avriel, "r-convex functions," Mathematical Programming 2 (1972), 309-323.

[3] M. Avriel, "Solution of certain nonlinear programs involving r-convex functions," J. Optimization Theory and Applications 11 (1973), 159-174.

[4] M. Avriel, "Nonlinear Programming: Analysis and Methods," (Prentice Hall, New Jersey, 1976).

[5] M. Avriel and I. Zang, "Generalized convex functions with applications to nonlinear programming," in: P. van Moeseke, ed., Mathematical Programs for Activity Analysis, North-Holland Publishing, New York, 1974, pp. 23-33.

[6] V.J. Bowman and T.C. Gleason, "A note on second order conditions for pseudo-convexity," Working paper 44-74-75, Graduate School of Industrial Administration, Carnegie-Mellon University, May 1975.

[7] R.W. Cottle, "Manifestation of the Schur complement," Linear Algebra and its Applications 8 (1974), 189-211.

[8] R.W. Cottle and J.A. Ferland, "On pseudo-convex functions of nonnegative variables," Mathematical Programming 1 (1971), 95-101.

[9] R.W. Cottle and J.A. Ferland, "Matrix-theoretic criteria for the quasi-convexity and pseudo-convexity of quadratic functions," Linear Algebra and its Applications 5 (1972), 123-136.

[10] J.A. Ferland, "Quasi-convex and pseudo-convex functions on solid convex sets," Technical Report 71-4, Operations Research House, Stanford University, April 1971.

[11] J.A. Ferland, "Maximal domains of quasi-convexity and pseudo-convexity for quadratic functions," Mathematical Programming 3 (1972), 178-192.

[12] J.A. Ferland, "Mathematical programming problems with quasi-convex objective functions," Mathematical Programming 3 (1972), 296-301.

[13] A.V. Fiacco, "Second order sufficient conditions for weak and strict constrained minima," SIAM J. Appl. Math. 16 (1968), 105-108.

[14] F.R. Gantmacher, "The Theory of Matrices," Vol. I,(Chelsea, New York, 1960).

[15] O.L. Mangasarian, "Pseudo-convex functions," SIAM J. Control 3 (1965), 281-290.

[16] O.L. Mangasarian, "Nonlinear Programming,"(McGraw-Hill, New York, 1969).

[17] B. Martos, "Nem-lineáris programozási módszerek hatóköre" (Power of nonlinear programming methods), A Magyar Tudományos Akadémia Közgazdaságtudományi Intézetének Közleményei, 20, Budapest 1966.

[18] B. Martos, "Subdefinite matrices and quadratic forms," SIAM J. Appl. Math. 17 (1969), 1215-1223.

[19] B. Martos, "Quadratic programming with a quasi-convex objective function," Operations Research 19 (1971) 82-97.

[20] B. Martos, "Nonlinear programming: theory and methods," (North Holland, Amsterdam-Oxford, 1975).

[21] P. Mereau and J.G. Paquet, "Second order conditions for pseudo-convex functions," SIAM J. Appl. Math. 27 (1974) 131-137.

[22] J. Ponstein, "Seven kinds of convexity," SIAM Review 9 (1967) 115-119.

[23] S. Schaible, "Beiträge zur quasikonvexen Programmierung," Ph.D. Dissertation, Universität Köln, 1971.

[24] S. Schaible, "Quasi-convex optimization in general real linear spaces," Zeitschrift für Operations Research 16 (1972) 205-213.

[25] S. Schaible, "Quasi-concave, strictly quasi-concave and pseudo-concave functions," in: R. Henn, H.P. Künzi and H. Schubert, eds. Methods of Operations Research 17 (Anton Hain, Meisenheim, 1973), pp. 308-316.

[26] S. Schaible, "Quasi-convexity and pseudo-convexity of cubic functions," Mathematical Programming 5 (1973) 243-247.

[27] S. Schaible, "Second order characterizations of pseudo-convex quadratic functions," Department of Operations Research, Stanford University, November 1975, to appear in Journal of Optimization Theory and Applications 12, April 1977.

[28] J. Stoer and C. Witzgall, "Convexity and Optimization in Finite Dimensions I,"(Springer-Verlag, Berlin-Heidelberg-New York, 1970).

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. ~~REPORT NUMBER~~	14 TR-76-12
2. GOVT ACCESSION NO.	
3. RECIPIENT'S CATALOG NUMBER	
4. ~~TITLE (and Subtitle)~~	6 Second Order Characterizations of Pseudo-convex Functions.
5. TYPE OF ~~REPORT & PERIOD COVERED~~	9 Technical Report,
6. ~~PERFORMING ORG. REPORT NUMBER~~	
7. ~~AUTHOR(s)~~	10 Mordecai Avriel ~~and~~ Siegfried Schaible
8. CONTRACT OR GRANT NUMBER(s)	N00014-75-C-0267
9. PERFORMING ORGANIZATION NAME AND ADDRESS	Department of Operations Research Stanford University Stanford, CA 94305
10. PROGRAM ELEMENT, PROJECT, TASK ~~AREA & WORK UNIT NUMBERS~~	16 NR-047-143
11. CONTROLLING OFFICE NAME AND ADDRESS	11 Operations Research Program Code 434 Office of Naval Research Arlington, VA 22217
12. ~~REPORT DATE~~	June [illegible]75 12 30p.
13. NUMBER OF ~~PAGES~~	28
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	
15. SECURITY CLASS. (of this report)	Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution ~~is unlimited.~~

15 N00014-75- [illegible]

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

[illegible]

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear Programming
Generalized Convexity
Pseudoconvex Functions
Second Order Conditions
Bordered Determinants

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Second order characterizations for (strictly) pseudoconvex functions are derived in terms of extended Hessians and bordered determinants. Additional results are presented for quadratic functions.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

402 166 4/5